ENERGY INNOVATION POLICY & TECHNOLOGY LLC
AMERICA'S POWER PLAN
www.energyinnovation.org
98 Battery Street; San Francisco, CA 94111
power@energyinnovation.org

METRICS FOR ENERGY EFFICIENCY: OPTIONS AND ADJUSTMENT MECHANSIMS

BY ROBBIE ORVIS, SONIA AGGARWAL, & MICHAEL O'BOYLE ● APRIL 2016

INTRODUCTION

In many states, efficiency targets have typically been achieved through programs requiring or incenting utilities to meet specific reductions in electricity use, often through complex calculations of measure-specific savings (i.e. efficiency technology, like an LED light bulb or high efficiency water heater). While measure-by-measure approaches seem to be common practice, some states, like California, have experienced significant problems with this approach.

As states consider fundamental changes in the utility revenue model, in particular linking utility performance to compensation, regulators will have to choose which metrics best track the outcomes they want from the electricity sector, including energy efficiency. This white paper makes the case that outcome-oriented, economy-wide metrics for energy efficiency are the most appropriate method for tying revenue to performance, and offers a menu of options to inform the process.

WHY CHOOSE AN OUTCOME-ORIENTED METRIC?

Utility energy efficiency programs in most states target specific energy reductions through evaluations of installed "measures," i.e. components such as light bulbs, boilers, and air conditioners. To estimate measure-by-measure savings, utilities, public utility commissions, and/or third parties perform complex calculations based on the number of installed measures and estimates of how much energy these measures save. These savings are then adjusted by a net-to-gross ratio, which adjusts savings based on an estimate of what would have happened without utility programs, the type and age of buildings in which the measure is installed, and other important factors.

While a measure-by-measure savings approach may be appropriate for some programs, it also introduces significant challenges, especially when revenue is tied directly to these savings estimates. Specifically, many of the assumptions that go into these savings estimates have "significant levels of uncertainty as well as annual variation."[1] When utility incentive payments are based on calculations with highly uncertain assumptions there is bound to be regulatory

---

[1] "Proposed Energy Efficiency Risk-Reward Incentive Mechanism and EM&V Activities." *California Public Utilities Commission, The Energy Division*. April 1, 2009.

conflict. One such example is the Risk-Reward Incentive Mechanism in California, where disagreements over the assumptions underlying savings estimates led to a failure of the program.[2]

As policymakers and regulators look to implement new efficiency incentive programs or improve existing programs, it is useful to consider alternatives to the traditional measure-by-measure savings approach.

A better alternative to a measure-by-measure estimation approach is an outcome-oriented metric, for example, total sales or kilowatt-hours (kWh) per customer. Outcome-oriented metrics look holistically at a utility's performance as reflected in the achievement of reduced energy consumption. As opposed to backward-looking accounting-based measurement, outcome-oriented metrics better reflect whether a utility has achieved the outcome or goal for system-wide efficiency set by regulators and enable utilities greater freedom in how they choose to pursue those outcomes.

Outcome-oriented metrics, if well-designed, may provide several benefits over traditional measure-by-measure programs. First, program performance may be significantly easier to monitor relative to measure-by-measure programs. For example, tracking a utility's total sales, or its sales per customer, is much more straightforward and transparent than requiring the utility to track installations and calculate savings for each of these installations. Second, outcome-oriented metrics can be more easily tied to state goals—for example, greenhouse gas targets—than measure-by-measure programs, which often rely on studies of potential rather than policy goals. Third, the administrative burden can be substantially reduced if commissions rely on an outcome-oriented metric rather than requiring detailed measure-by-measure analysis to determine utility performance.[3]

IMPORTANT CONSIDERATIONS FOR OUTCOME-ORIENTED METRICS

There are several important criteria to consider when developing outcome-oriented metrics. The following principles should guide regulators and utilities considering measuring performance that is tied to outcomes:

1. Tie metrics to broad policy targets
2. Minimize the need for ex post adjustment mechanisms

---

[2] See: Orvis, R., "Lessons for Designing Counterfactuals in Earnings Incentive Mechanisms: California as a Case Study," *Energy Innovation LLC.* February, 2015.

[3] California's experience with the RRIM and measure-by-measure estimates demonstrates that overly burdensome evaluation, measurement, and verification (EM&V) can be very costly. In terms of budgeting, for the 2006-2008 program cycle, the CPUC authorized $163 million in spending for evaluation, measurement, and verification (EM&V). The EM&V funding amounted to 7.6 percent of funding for the state's whole efficiency portfolio spending, which, relative to a U.S. average of 3 percent, is extraordinarily high—two to three times greater than other states. See: "2014 State of the Efficiency Program Industry: Budgets, Expenditures, and Impacts" *Consortium for Energy Efficiency*. May 1, 2015.

3. Maximize transparency and predictability of adjustment mechanisms, if needed

TIE METRICS TO BROAD POLICY TARGETS

One of the advantages of using outcome-oriented metrics is the ability to directly tie the metrics to broader policy goals. For example, as discussed above, measure-by-measure efficiency programs often estimate savings goals based on studies of market potential, rather than on policy targets. Studies of future potential often take a bottom-up approach, which can underestimate achievable or economic efficiency savings.[4] Conversely, outcome-oriented metrics tied to statewide targets can push utilities to think creatively about how to meet those targets rather than being limited to the measures determined in studies.

Similarly, targets can be logically related to one another to collectively drive towards several consonant policy goals. For example, a CO_2 per MWh standard for generation could be combined with a kWh per capita standard for efficiency. Together, they form a CO_2 per capita value, which can be tied to overall policy-driven greenhouse gas targets.[5]

MINIMIZE THE NEED FOR ADJUSTMENT MECHANISMS

Adjustment mechanisms are formulas or methods to account for exceptional events beyond utilities' control that materially affect performance on a given metric. Adjustment mechanisms are typically used when their absence would result in arbitrary performance evaluations, for example, record heat causing a sustained spike in demand.

Taking these concerns into account, metrics should be designed with the goal of minimizing the need for adjustment mechanisms. Minimizing adjustment allows metrics to target decisions and outcomes that are primarily under the utility's control. However, outcome-oriented metrics are likely to require some form of adjustment because at least some part of the metric is very likely to be influenced by events far outside the utility's control. For example, weather, the economy, and population growth are all factors that can affect outcomes. As discussed later below, there are ways to simplify the adjustment process when adjustment is necessary.

A large degree of conflict can be avoided if metrics can be streamlined with minimal need for adjustment. Adjustment can be built into metrics. For example, certain metrics such as kWh per unit of GDP, kWh per customer, or CO_2 per customer account for population or economic changes within the metric itself. Metrics can also be running averages (e.g. three to five years) to account for annual variations in weather or economic trends. The less adjustment takes place after measurement, the smaller the disagreements between adversarial parties, and the greater the regulatory efficiency.

[4] See Orvis et al., "Accelerating Carbon Reductions from California's Electricity Sector (p. 19-20)." *Energy Innovation LLC.* March 2015. (http://energyinnovation.org/wp-content/uploads/2015/03/2015_CA_ElectricityPolicy_FINAL.pdf)

[5] Ibid.

MAXIMIZE TRANSPARENCY AND PREDICTABILITY OF ADJUSTMENT MECHANISMS

To the extent that adjustment mechanisms are required, efforts should be taken to maximize the transparency of these mechanisms and to set clear rules ahead of time for how these mechanisms will be used, so that there is less room for an ex post dispute.[6] For example, agreeing on data sources ahead of time and relying on indisputable data like the number of customers, officially published weather data, or officially published economic data, is bound to be less contentious than relying on empirical or counterfactual studies over which there is a high potential for disagreement (for example, estimates of free-ridership).

When revenue is at stake, the utility and regulators have adverse incentives for adjustment. Even in cases where assumptions may be relatively straightforward, utilities are likely to challenge the adjustment if doing so may result a significant increase in revenue.

At the same time, regulators will tend to revise the adjustment if the financial incentive doesn't comport with their expectations. To avoid potential disputes, adjustment mechanisms should adjust utility metrics based on independently verified data, and not present an opportunity to re-evaluate the metric itself.

POTENTIAL OUTCOME-ORIENTED METRICS

There are many outcome-oriented metrics that can be used to track energy efficiency, ranging from high-level (e.g. total kWh sales) to very nuanced (e.g. kWh per ton cement clinker production). Outcome-oriented metrics generally fall into two categories: energy intensity metrics and energy consumption metrics. Intensity metrics measure the energy use on a per unit level, for example, per capita, per customer, or per household. Consumption metrics measure overall energy use, for example, energy sales or percent overall improvement.

The table below introduces seven options for outcome-oriented metrics that may serve as a starting point for tracking utility energy efficiency, each of which may be tailored to more specific circumstances.

Table 1: Seven Options for Outcome-Oriented Metrics for Energy Efficiency

Metric	Type	Description
kWh per capita	Intensity	Measure energy use per person based on retail sales and estimates/projections of a utility territory's population.
kWh per customer	Intensity	Measure energy use per customer (i.e. meter) based on retail sales and a count of meters/customers in a utility's territory.

[6] See, e.g., R. Orvis memo on Counterfactuals, note 2 above.

kWh per household	Intensity	For residential sector: Measure energy use per household based on retail sales and an estimate of the number of households in a utility's territory
kWh per square foot	Intensity	For commercial sector: Measure energy per square foot (or other unit area) based on retail sales and an estimate of square footage of building space within a utility's territory.
kWh per $ GDP	Intensity	Measure energy use per dollar GDP based on retail sales and estimate of GDP within a utility's service territory.
Total sales (kWh)	Consumption	Measure total retail sales by a utility.
Annual percent improvement from start year or baseline year	Consumption	Track the annual percent reduction in sales by a utility. Note: this is intended to imply a real-world measurement, not a reduction from a counterfactual.

ADJUSTMENT MECHANISMS

Outcome-oriented metrics and targets are likely to need adjustment to account for effects outside of the utility's control, especially if a significant amount of revenue is tied to performance. Weather, the economy, and changes in the number of customers or persons within a utility territory are likely to be the largest sources of exogenous effects on metrics and will likely require adjustment mechanisms. This section examines best practices from other regions that show what is possible, but it is also important to keep in mind that many of these adjustments can be avoided or minimized by using the right metric.

ADJUSTING FOR WEATHER

A good deal of experience with adjusting utility revenue for weather has been achieved through decoupling programs across the U.S., though these adjustment programs have primarily focused on natural gas utilities. As of 2013, 14 utilities had gas decoupling with weather adjustment mechanisms.[7] It is worth noting that the majority of *electricity* decoupling programs do not adjust for weather. This likely reflects the fact that weather fluctuations often affect natural gas use more than electricity use, as well as a desire on the part of regulators to keep the risk of weather on utilities rather than shifting this risk to customers. But if regulators add a symmetrical performance incentive and penalty, additional risk is shifted to the utility and it may be prudent to remove exogenous effects in order to engender confidence on the part of the utility that its decisions and investments will directly impact the performance metric.

[7] Pamela Morgan. "A Decade of Decoupling for US Energy Utilities: Rate Impacts, Designs, and Observations (p. 12)" *Graceful Systems LLC*. February 2013.

Natural gas weather normalization provides a useful starting point for evaluating weather normalization approaches as many utilities calibrate revenue adjustments based on observed weather data.

One relevant example is Southwest Gas (SWG) in Arizona. SWG uses a two-part mechanism to determine adjustments to the recovery of fixed costs on customer bills;[8] the first component adjusts volume based on weather deviations from historical temperatures, and the second component adjusts volume based on changes in customer behavior. SWG uses the lower of the two components to normalize sales against weather.

The first approach is a "billing cycle analysis volume adjustment." SWG takes the difference in actual (observed) heating degree days (HDDs)[9] and normal (a typical year) HDDs to estimate the increased number of days of heating use in a billing cycle. Next, SWG determines the average use per HDD per customer based on a customer's metered use (less the base amount) and the number of HDDs in a billing cycle. When multiplied together, this results in a consumption adjustment that is applied to customer bills for collection of fixed charges.[10]

The second approach is a "multi-season analysis volume adjustment," which uses data from the previous 24 months to develop average monthly values and adjust again based on HDDs using a model. Under this approach SWG uses a regression model to compare a customer's historical monthly use to actual weather in each billing cycle and then to adjust this amount by the observed variance around use per HDD for the current. The advantage of this approach is that it accounts for the variability of a customer's use on average, as opposed to the previous approach, which does not account for variation in customer behavior.

By using these approaches in combination, these programs not only keep utilities from over-recovering during demand drops due to weather, but also protect them from downside risk of energy efficiency investments. This "limited" or "partial" decoupling is a way of compensating utilities for efficiency investments while not "compensating" them for weather.[11]

---

[8] "Decoupling and Energy Efficiency Enabling Provision (EEP)." *Southwest Gas Corporation.* (www.swgas.com/en/decoupling-and-eep).

[9] Heating and cooling degree days are defined as the difference between the average daily temperature and threshold temperature for heating and cooling. NOAA uses a threshold of 65 for both heating and cooling. See: http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/cdus/degree_days/ddayexp.shtml for more information.

[10] E.g., there are 25 HDDs in December relative to a "normal" of 20 HDDs. Over those 25 HDDs a customer used 50 Therms, or 2 Therms/HDD. Since the revenue recovery for fixed costs is a *rate* based on an estimated 20 HDDs, the customer is now overpaying the utility. The total recovery for fixed costs is adjusted to be based on 20 HDD * 2 Therms/HDD = 40 Therms, multiplied by the fixed cost recovery rate.

[11] Other utilities use similar approaches of calibrating energy use based on heating and cooling degree days. See for example: http://www.centerpointenergy.com/en-us/Corp/Documents/Minnesota%20Rates%20and%20Rights//Section5-ConservationEnablingRider.pdf.

For the purposes of calibrating an energy efficiency metric for electric utilities, regulators should consider a two-step approach. First, a study should be conducted evaluating the extent to which weather (heating and cooling degree days) affects energy use. Air conditioning use in the summer is likely to be correlated with cooling degree days. To the extent that customers use electric heating during the winter, electricity use may be correlated with heating degree days. The same lessons could be applied to winter and summer peak reduction targets as well.

States that experience a significant amount of load variation due to weather should move forward with the second step of developing adjustment mechanisms using the examples above as a starting point. Adjustments could be made by customer class to increase the likelihood of finding a statistically significant relationship between weather and load (i.e. customer classes are also likely a significant determinant of load variability, so accounting for this is important).

During this early stage of tying compensation to utility performance, regulators might also consider starting with measuring multiple metrics to understand which best account for weather-based variability and obviate the need for adjustment.

ADJUSTING FOR THE ECONOMY

Economic factors can play an important role in utility sales, and may need to be accounted for in an outcome-oriented utility metric. The most straightforward indicator which can explain some variation is GDP. The best places to find information relating GDP to electricity consumption are in a utility's integrated resource plan (IRP) or its rate case.

The rate-making process requires utilities to develop an estimate of forecast load in order to apportion its revenue requirement to customers and develop rates. In developing load estimates, utilities should (and most already do) include estimates of the effect of GDP on electricity use. Similarly, utilities may develop multiple scenarios that make different assumptions about economic growth. For example, in its 2015 IRP, Idaho Power developed a low, expected, and high load growth rate reflecting different assumptions about economic and demographic conditions.[12] Consolidated Edison provided similar low, medium, and high scenarios for gross energy and peak demand growth in its 2010 Long Range Plan.[13] Regulators could use similar information from utility resource plans and rate cases to adjust a metric based on fluctuations in GDP.

ADJUSTING FOR POPULATION

Utilities may be concerned about adjusting a metric based on the number of customers in a utility's service territory. There are a couple of ways to approach this issue. First, intensity metrics can avoid this issue altogether. For example, a kWh per customer standard accounts for the number of customers in the standard itself, which gets around this issue.

---

[12] "Integrated Resource Plan 2015: Appendix A Sales and Load Forecast." *Idaho Power Company*. June 2015.

[13] "Electric System Long Range Lan: Appendix A, Demand and Energy Forecasts (p. 5-27)." *Consolidated Edison.* 2010 (http://www.coned.com/publicissues/PDF/ESLRP%20Appendices%20December%202010%20Final.pdf).

Another option is to "freeze" the customer base whose meters were used in determining the metric (which assumes projected load growth), and track this cohort over time to determine the value of the metric. With the denominator fixed, utility performance on efficiency is the only lever for meeting the target. This ensures that there is no disincentive to acquire new customers, particularly large industrial or commercial users that bring jobs to communities, but still fairly tracks utility performance in the service territory.

ADJUSTING FOR ATTRIBUTION

Many efficiency programs focus heavily on the "attribution problem," or determining what proportion of installed measures and/or savings are due exclusively to the utility's intervention and what proportion are due to exogenous factors. While it is important not to overcompensate the utility, especially for actions it doesn't actually take, too much focus on the attribution problem can distract the regulatory process from actually achieving energy savings. In California, arguments over the net-to-gross ratio—the variable used to address attribution—became a focus of the shareholder incentive mechanism that inhibited the ability of the program to drive energy savings. Furthermore, attribution calculations are inherently prone to conflict because they rely to a large degree on highly uncertain parameters.[14] Finally, attribution parameters tend to differ significantly across different efficiency measures, and developing a utility-wide attribution estimate may prove very difficult.

An important consideration when sizing a utility incentive is the amount of risk involved in making investment decisions with uncertain awards. For example, when regulators in California moved to allow bonus payments based on ex ante savings claims as opposed to ex post verification, they lowered the utility's rate of return to reflect the fact that the utility had a lower risk of its award payments deviating from expectations.[15] Other states can address the attribution issue by lowering the incentive award on efficiency; if utilities will get compensated for non-utility investments, then they should earn less of a return on a metric measuring overall reductions.

Regulators may also increase the ambition of a metric to account for non-utility influence when market forces are already driving improved performance. This is especially important for energy efficiency, which has significant uptake regardless of utility programs because of third-party marketers and savvy customers.

Outcome-oriented metrics, by their nature, focus less on attribution than measure-by-measure approaches to energy efficiency. This is preferable for many reasons, but regulators should pay attention to the effect of metric design on where risks are borne between customers and the

---

[14] Orvis, R. "Lessons for Designing Counterfactuals in Earnings Incentive Mechanisms: California as a Case Study" *Energy Innovation LLC.* February 2015.

[15] "Decision Adopting Efficiency Savings and Performance Incentive Mechanism." *California Public Utilities Commission*. 2013.

utility. To the extent utility risk is increased by focusing on outcomes, utility upside should also increase. Likewise, to the extent risks are mitigated, utility upside should be limited.

OTHER OPTIONS

In addition to the steps mentioned above, there are other simple steps that can be taken to normalize efficiency metrics. One option is to use a running average (e.g. every three years or five years). This way, the metric captures the long-term trend in energy efficiency, but is less susceptible to annual variations and swings in load based on weather, the economy, and other factors.

Another approach is to filter out data at the tail-ends of daily energy use over a given area. For example, in calculating the metric, utility regulators could allow utilities to remove the five or 10 percent of days with the highest and lowest electricity use. This has the effect of eliminating outliers, but still captures the underlying trend and utility performance.

Finally, while regulators should establish targets far enough into the future to provide investment certainty, a periodic review (e.g. once every 3-5 years) of the metric and a bounded (i.e. max/min) adjustment factor can help normalize the target over time.

RECOMMENDATIONS FOR UTILITY REGULATORS

There are many energy efficiency metrics for utilities and regulators to choose from, a few of which are discussed above. Some of these metrics have clear advantages over one another when considering measuring performance and tying it to compensation. The best metrics are those that are easy to measure, require minimal adjustment, and have transparent adjustment mechanisms where adjustment is necessary.

The kWh per customer, total sales, and annual percentage improvement metrics are most likely the easiest metrics to measure and the hardest for the utility to manipulate. The first of these, kWh per customer, is calculated easily by dividing the sales in each customer class by the number of customers in each customer class. This information is readily available as part of the rate case, so this should be a relatively straightforward metric. This metric can be also be normalized by only including consumption from meters present when measurement began. Total sales is the easiest consumption metric to measure, while percentage improvement is also easy as it is simply a function of total sales over multiple years.

Some of the other metrics may prove more challenging to track, but incorporate normalization in the metric. KWh per capita and kWh per household are similar to kWh per customer in that they are naturally normalized for population growth, but require an additional estimate of the number of residents or households *within* a utility's territory. This may be complicated by the fact that utility territories do not tend to overlap cleanly with census measurement areas (e.g. census tracts or blocks). Thus, utilities and regulators will have to agree on population and household estimates and apportion these to utility territories, which opens an avenue to confrontation. Similarly, it is not clear how these standards would operate for non-residential customers.

The kWh per GDP metric suffers from a similar setback. While the metric is desirable because it inherently accounts for changes in the economy, it requires utilities and/or regulators to develop GDP estimates particular to specific utility territories (and possibly even to customer classes). It may also create a strange and undesirable incentive to limit economic growth if it would adversely affect the efficiency metric. A statewide GDP value could be used if empirical research demonstrated that GDP changes within the state and across customer classes are relatively homogenous or have a similar effect.

The kWh per square foot metric is also intriguing since it attempts to control for a major component of energy demand—building area. While this metric seems promising, there is limited data available on the square footage of buildings, especially within the confined geographic area of a utility's service territory. If utilities have estimates or documentation of square footage, this could be a useful metric, but such estimates must overcome the lack of transparency and information asymmetry that would lead to regulatory conflict.

Based on its ease of measurement and fewer adjustment requirements, a kWh per customer metric may be the best metric of those discussed in this memo for several reasons; it is easy to measure, transparent, and hard to game. The number of customers is an easily quantifiable value (i.e. the number of meters) and the metric can be calibrated to avoid counting new customers and biasing the metric, which is more difficult with the other metrics. The kWh per customer metric also reduces the potential for adversarial proceedings between the utility and the commission because it only requires adjustment for weather, which can be done based on the description above.[16] Additionally, it is an outcome-oriented metric focused less on attribution and can be linked easily to the outcomes that maximize value in the electricity system.

Regardless of which metric is ultimately selected, the design of these metrics should strive to limit the need for adjustment and, where necessary, adjust based on transparent mechanisms and empirical research.

[16] Economic growth is more or less captured by keeping the number of meters fixed in the metric. For energy consumption due to individual personal economic growth, regulators may feel that risk should fall on utilities, who can facilitate low-energy alternatives that still promote economic growth.